UC Irvine
ICS Technical Reports

Title

Heterogeneous arrival streams, burstiness and packet discarding : a study of individual packet loss

Permalink

https://escholarship.org/uc/item/4n50b8mz

Authors

Bae, Jaime Jungok
Suda, Tatsuya
Simha, Rahul

Publication Date

1991

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Notice: This Material
may be protected
by Copyright Law
(Title 17 U.S.C.)

Z
699
C3
no. 91-58

LIBRARY
University of California
IRVINE

Heterogeneous Arrival Streams, Burstiness and Packet Discarding: A Study of Individual Packet Loss

Jaime Jungok Bae, Tatsuya Suda and Rahul Simha
Technical Report No. 91-58

Department of Information and Computer Science
University of California, Irvine
Irvine, CA 92717

Analysis of a Finite Buffer Queue with Heterogeneous Markov Modulated Arrival Processes: A Study of the Effects of Traffic Burstiness on Individual Packet Loss

JAIME JUNGOK BAE AND TATSUYA SUDA

Department of Information and Computer Science
University of California, Irvine
Irvine, CA 92717
(phone) 714-856-5474, (fax) 714-856-4056
(e-mail) jbae@ics.uci.edu, suda@ics.uci.edu

RAHUL SIMHA

Department of Computer Science
College of William and Mary
Williamsburg, VA 23185
(e-mail) simha@cs.wm.edu

Abstract

This paper considers a queueing system with a finite buffer and multiple heterogeneous arrival streams. We focus on Markov Modulated Arrival Processes with differing burstiness and investigate the loss of individual arrival streams when the parameters of the heterogeneous arrival streams are varied.

Our analysis follows a new stochastic integral approach and extends existing results for continuous-time Markov Modulated Poisson Arrival Processes to include both continuous-time and discrete-time treatments of multiplexed heterogeneous Markov Modulated Arrivals. In addition, we derive loss probabilities for a priority packet discarding scheme (a congestion control mechanism suitable for high-speed networks). Several interesting numerical results are presented; we introduce a new characterization of an arrival stream, which we refer to as self-loss, and use this to qualitatively predict the effects of multiplexing bursty streams with non-bursty streams.

1 Introduction

B-ISDN and high-speed networks are expected to support diverse applications such as voice, moving image, data transfer and teleconferencing. Different applications have different traffic characteristics and require different grades of service (GOS). Many architectures proposed for high-speed networks, such as Asynchronous Transfer Mode (ATM) and IBM's PARIS [5] are based on packet switching and explicitly permit packet loss in order to gain bandwidth efficiency. Therefore, it is important to predict whether a network can provide a required grade of service (i.e., an acceptable level of packet loss) for each of the service classes on a network.

In this paper, we study a queueing system where heterogeneous traffic streams (sessions) are multiplexed and investigate how the heterogeneity of the arrival streams, especially the varying level of burstiness, affects packet loss in individual streams. We consider the class of Markov Modulated Arrival (MMA) streams both in continuous time (a Markov Modulated Poisson Process or MMPP) [10] and in discrete-time (a Markov Modulated Bernoulli Process or MMBP) and present an exact analysis of individual packet loss for MMA streams.

We also analyze packet loss when a packet discarding control scheme is applied as a congestion control mechanism. Several buffer control schemes have been proposed to alleviate the problem of packet loss. For example, low priority packets may be discarded when the buffer is filled to a certain level. In these cases it is especially important to predict packet losses for individual streams with different priorities. We study a previously proposed packet discarding scheme [3] and derive expressions for individual packet loss.

An important contribution of this paper is the study of the impact of burstiness of traffic streams on the individual packet loss probabilities. We present several numerical results describing individual packet loss probabilities when bursty streams are multiplexed with non-bursty streams. We introduce the concept of self-loss for a single stream, the packet loss incurred when a stream is multiplexed with itself, and show how the self-loss of bursty and non-bursty streams may be used to understand the effects of multiplexing heterogeneous arrival streams on individual packet loss.

Most analytical approaches in the past study the packet loss incurred when several identical arrival streams are multiplexed at a single buffer [1, 10, 16], and thus fail to adequately address the issue of the heterogeneity of the arrival streams. In this paper, we obtain individual packet loss for both continuous and discrete-time cases, as well as when a buffer control scheme (a packet discarding scheme) is in effect. We follow the stochastic integral approach in [19], a method of independent interest, to derive our individual packet loss expressions. In doing so, we re-derive expressions for the continuous-time case presented without proof in [15]. In [15], the emphasis was on analyzing parcel overflow processes using a two-state MMPP approximation for a multistate MMPP. Numerical results presented therein focused on the accuracy of the approximation.

The rest of this paper is organized as follows. In sections 2 and 3, individual packet loss probabilities are derived for continuous and discrete-time cases when two arrival streams are present. In Section 4, our discrete-time analysis is extended to accommodate a priority packet discarding scheme. In Section 5, our analysis is applied to investigate the effects of individual traffic characteristics and traffic mix on the individual packet loss probabilities; we present several interesting numerical results in this section. Finally, in Section 6, we make some concluding remarks and discuss future work. Note that our analyses for 2-stream cases can easily be extended for $N(>2)$ heterogeneous traffic streams. In Appendices A and B, analysis for $N(>2)$ heterogeneous streams is presented for continuous and discrete time, respectively. In Appendix C, the analysis of a priority packet discarding scheme for 2-stream case is extended to $N(>2)$ heterogeneous streams.

2 Continuous-Time Case

In this section, individual packet loss probabilities are obtained for the continuous-time case. Each arrival stream is modeled by a 2-state MMPP. Note that a 2-state MMPP is a fairly general process – by selecting appropriate parameter values, a 2-state MMPP can represent a Poisson process (suitable to describe data arrivals) and an Interrupted Poisson Process (suitable to describe On/Off traffic sources such as voice). 2-state MMPPs have been used to represent a superposition of several identical sources [10] and thus, each arrival stream in our model can be viewed as a single source or a superposition of multiple identical sources. In this section, for simplicity, it is assumed that two heterogeneous streams (stream A and stream B) are multiplexed. The analysis can easily be extended to a case where $N(>2)$ heterogeneous input streams are multiplexed and is discussed in Appendix A.

2.1 Model and Notations

Consider a single first-come-first-served queue driven by two 2-state MMPP arrival processes. The queue has a finite buffer space whose maximum size is $K-1$ (packets). Thus, the maximum system size (the maximum buffer size plus the packet being served) is K (packets). Service times of packets from streams A and B are exponentially distributed with rate μ. Packets from each of streams A and B arrive according to a 2-state MMPP. A 2-state MMPP is characterized by two alternating 'driving' states. It is usually assumed that the duration of each state is exponentially distributed and packet arrivals in each state are Poisson processes with different rates.

The driving states of stream A are labeled 1 and 2; the driving states of stream B are labeled 3 and 4. (Refer to Figure 1.) For stream A, the transition rate from state 1 to state 2 is denoted by α, and the transition rate from state 2 to state 1 is denoted by β. For stream B, the transition rate from state 3 to state 4 is denoted by γ, and the transition rate from state 4 to state 3 is denoted by

δ. Thus, for stream A, the sojourn times in states 1 and 2 are exponentially distributed with the mean $1/\alpha$ and $1/\beta$, respectively. For stream B, the sojourn times in states 3 and 4 are exponentially distributed with the mean $1/\gamma$ and $1/\delta$, respectively. The generation of packets when the MMPP is in state i follows a Poisson process with rate λ_i. Thus, when stream A is in state i ($i = 1, 2$) and stream B is in state j ($j = 3, 4$), the aggregate arrival rate to the queueing system is $\lambda_i + \lambda_j$.

Define $Y_A(t)$ and $Y_B(t)$ as the states of stream A and B MMPP's at time t, respectively, i.e., $Y_A(t) = 1$ or 2, and $Y_B(t) = 3$ or 4. For $i = 1, 2$ and $j = 3, 4$, we define the following indicator functions:

$$I_i(t) = \begin{cases} 1, & \text{if } Y_A(t) = i \\ 0, & \text{otherwise} \end{cases}$$

$$I_j(t) = \begin{cases} 1, & \text{if } Y_B(t) = j \\ 0, & \text{otherwise.} \end{cases}$$

$I_i(t)$ ($I_j(t)$) becomes 1, if the state of the MMPP for stream A (B) at time t is i (j). Otherwise it is 0.

Let $Z(t)$ ($0 \leq Z(t) \leq K$) denote the system state (i.e., the number of packets in the system including both a server and a buffer) at time t. Define the following indicator function for a system state q ($0 \leq q \leq K$).

$$U_q(t) = \begin{cases} 1, & \text{if } Z(t-) = q \\ 0, & \text{otherwise.} \end{cases}$$

$U_q(t)$ is 1, if the system state at time $t-$ is q. Otherwise it is 0.

Let $N_A(t)$ and $N_B(t)$ be the cumulative number of arrivals from stream A and from stream B in the time interval $[0, t]$, respectively. Let $N(t)$ be the cumulative number of arrivals in the time interval $[0, t]$. Thus,

$$N(t) = N_A(t) + N_B(t). \tag{1}$$

Let $\Lambda_A(t)$ and $\Lambda_B(t)$ denote the *compensators* [4] for the processes $N_A(t)$ and $N_B(t)$, respectively, so that $N_A(t) - \Lambda_A(t)$ and $N_B(t) - \Lambda_B(t)$ are martingales. (See [14], pp.239, for the definition of compensators, and see [4], pp.4, for the definition of martingales.) For instance, for the Poisson process, we have the compensator $\Lambda(t) = \lambda t = \int_0^t \lambda ds$, and for the doubly stochastic Poisson process, we have the compensator $\Lambda(t) = \int_0^t \lambda(s) ds$. (See [4]). For our model, the intensity function $\lambda_A(t)$ for stream A and $\lambda_B(t)$ for stream B are given by

$$\lambda_A(t) = I_1(s)\lambda_1 + I_2(s)\lambda_2, \qquad \text{and} \qquad \lambda_B(t) = I_3(s)\lambda_3 + I_4(s)\lambda_4. \tag{2}$$

Thus, the compensators for $N_A(t)$ and $N_B(t)$ become

$$\Lambda_A(t) = \int_0^t (I_1(s)\lambda_1 + I_2(s)\lambda_2) ds, \qquad \text{and} \qquad \Lambda_B(t) = \int_0^t (I_3(s)\lambda_3 + I_4(s)\lambda_4) ds. \tag{3}$$

We define the following limiting probabilities. Let $\pi(i,j,q)$ ($i=1,2, j=3,4, 0 \le q \le K$) be the limiting distribution for the Markov process $\{Y_A(t), Y_B(t), Z(t)\}$. Let $\pi(i,q)$ ($i=1,2, 0 \le q \le K$) be the limiting distribution for the Markov process $\{Y_A(t), Z(t)\}$, and $\pi(j,q)$ ($j=3,4, 0 \le q \le K$) be the limiting distribution for the Markov process $\{Y_B(t), Z(t)\}$. Note that $\sum_j \pi(i,j,q) = \pi(i,q)$ and $\sum_i \pi(i,j,q) = \pi(j,q)$.

2.2 Analysis

In this analysis, we obtain the following probabilities:

1. the long term probability $P_A(q)$ that an arrival from stream A sees the system in state q,
2. the long term probability $P_B(q)$ that an arrival from stream B sees the system in state q, and
3. the long term probability $P(q)$ that an arbitrary arrival sees the system in state q.

Note that $0 \le q \le K$. From these probabilities, we can easily obtain the loss proabilities for stream A ($P_{loss}(A)$) and for stream B ($P_{loss}(B)$) by the following:

$$P_{loss}(A) = P_A(K) \quad \text{and} \quad P_{loss}(B) = P_B(K). \tag{4}$$

Further, the overall packet loss probability $P_{loss}(O)$ for the aggregated arrival process (i.e., the loss probability of packets, indistinguishing streams A and B) is given by

$$P_{loss}(O) = P(K) \tag{5}$$

First, let us calculate the long term probability $P_A(q)$ for an arrival from stream A to see the system state q. We have

$$P_A(q) = \lim_{t\to\infty} \frac{1}{N_A(t)} \int_0^t U_q(s)dN_A(s) = \lim_{t\to\infty} \frac{t}{N_A(t)} \lim_{t\to\infty} \frac{1}{t} \int_0^t U_q(s)dN_A(s). \tag{6}$$

By noting that the term $\lim_{t\to\infty} \frac{N_A(t)}{t}$ (the inverse of the first term on the right hand side of the above equation) gives the mean arrival rate of stream A, we obtain the following expression:

$$\lim_{t\to\infty} \frac{N_A(t)}{t} = \frac{\lambda_1 \frac{1}{\alpha} + \lambda_2 \frac{1}{\beta}}{\frac{1}{\alpha} + \frac{1}{\beta}} = \frac{\lambda_1\beta + \lambda_2\alpha}{\alpha+\beta}. \tag{7}$$

To obtain the term $\lim_{t\to\infty} \frac{1}{t} \int_0^t U_q(s)dN_A(s)$ in Eq.(6), we use the following:

1. The intensity function $\lambda_A(t)$ for stream A is bounded (i.e., $||\lambda_A(t)|| = \sup_{t\ge0} |\lambda_A(t)| < \infty$.)

2. $U_q(t)$ is a predictable process.

3. $E(\int_0^t |U_q(s)|d|\Lambda_A(s)|) < \infty$, for every $t \in (0, \infty)$.

The three assertions above are easily verified. Note that $N_A(t)$ is, by definition, a doubly stochastic Poisson process [17]. $\lambda_A(t)(= I_1(t)\lambda_1 + I_2(t)\lambda_2)$ is bounded since it is equal to either λ_1 or λ_2 which are finite. $U_q(t)$ is a predicatable process since it is left continuous. (Proof of this is given in [4], pp.9.) The last condition may be shown as follows:

$$\begin{aligned} E(\int_0^t |U_q(s)|d|\Lambda_A(s)|) &\leq E(\int_0^t d|\Lambda_A(s)|) = E(\int_0^t |(I_1(s)\lambda_1 + I_2(s)\lambda_2|ds) \\ &\leq E(\int_0^t \max(\lambda_1, \lambda_2)ds) \\ &= E(\max(\lambda_1, \lambda_2)s|_0^t) = \max(\lambda_1, \lambda_2) \times t. \end{aligned}$$

Since both λ_1 and λ_2 are finite,

$$E(\int_0^t |U_q(s)|d|\Lambda_A(s)|) < \infty.$$

Now, we use the following theorem [19] to obtain the term $\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN_A(s))$ in Eq.(6).

Theorem 1: Let $T = [0, \infty)$. For $t \in T$, assume $N_A(t)$ is a doubly stochastic Poisson process with bounded intensity function $\lambda_A(t)$. Define $R(t)$ as the following:

$$R(t) = \int_0^t U_q(s)dN_A(s) - \int_0^t U_q(s)d\Lambda_A(s),$$

where $U_q(t)$ is the indicator function for a system state q and $\Lambda_A(t)$ is a compensator for $N_A(t)$. If $U_q(t)$ is a predictable process satisfying the following condition for every $t > 0$,

$$E(\int_0^t |U_q(s)|d|\Lambda_A(s)|) < \infty$$

then the following equation holds with probability one.

$$\lim_{t\to\infty} \frac{R(t)}{t} = 0 \tag{8}$$

For a proof of this theorem, see [19].

In [19], it is shown that the stochastic integral $R(t)$ is a martingale; intuitively, $R(t)$ 'hovers' around zero and thus $\lim_{t\to\infty} \frac{R(t)}{t} = 0$. We now apply Theorem 1 to obtain the term

$\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN_A(s)$ in Eq.(6). We have

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN_A(s) &= \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)d\Lambda_A(s) \\
&= \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)(I_1(s)\lambda_1 + I_2(s)\lambda_2)ds \\
&= \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)I_1(s)\lambda_1 ds + \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)I_2(s)\lambda_2 ds \\
&= \lambda_1\pi(1,q) + \lambda_2\pi(2,q). \qquad (9)
\end{aligned}$$

For the last step, we used the fact that $\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)I_1(s)ds$ ($\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)I_2(s)ds$) represents the limiting probability that $\{Y_A(t) = 1, Z(t) = q\}$ ($\{Y_A(t) = 2, Z(t) = q\}$), and that it is equal to $\pi(1,q)$ ($\pi(2,q)$). From Eqs.(6), (7) and (9), we have

$$P_A(q) = \frac{(\alpha+\beta)(\lambda_1\pi(1,q) + \lambda_2\pi(2,q))}{\lambda_1\beta + \lambda_2\alpha}. \qquad (10)$$

Using the same argument for stream B, we obtain

$$P_B(q) = \frac{(\gamma+\delta)(\lambda_3\pi(3,q) + \lambda_4\pi(4,q))}{\lambda_3\delta + \lambda_4\gamma}. \qquad (11)$$

Next, we compute the long term probability $P(q)$ of an arbitrary arrival seeing the system state q:

$$\begin{aligned}
P(q) &= \lim_{t\to\infty} \frac{1}{N(t)}\int_0^t U_q(s)dN(s) \\
&= \lim_{t\to\infty} \frac{t}{N(t)} \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN(s) \qquad (12)
\end{aligned}$$

Since $N(t) = N_A(t) + N_B(t)$, we obtain

$$\begin{aligned}
\lim_{t\to\infty} \frac{N(t)}{t} &= \lim_{t\to\infty} \frac{N_A(t)}{t} + \lim_{t\to\infty} \frac{N_B(t)}{t} \\
&= \frac{\lambda_1\beta + \lambda_2\alpha}{\alpha+\beta} + \frac{\lambda_3\delta + \lambda_4\gamma}{\gamma+\delta} \qquad (13)
\end{aligned}$$

For the last step, we used Eq.(7). For the second term in the right hand side of Eq.(12), we have

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN(s) &= \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)d(N_A(s) + N_B(s)) \\
&= \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN_A(s) + \lim_{t\to\infty} \frac{1}{t}\int_0^t U_q(s)dN_B(s) \\
&= \lambda_1\pi(1,q) + \lambda_2\pi(2,q) + \lambda_3\pi(3,q) + \lambda_4\pi(4,q) \qquad (14)
\end{aligned}$$

By substituting Eqs.(13) and (14) into Eq.(12), we have

$$P(q) = \frac{(\alpha+\beta)(\gamma+\delta)(\lambda_1\pi(1,q)+\lambda_2\pi(2,q)+\lambda_3\pi(3,q)+\lambda_4\pi(4,q))}{(\lambda_1\beta+\lambda_2\alpha)(\gamma+\delta)+(\lambda_3\delta+\lambda_4\gamma)(\alpha+\beta)}. \tag{15}$$

In order to obtain probabilities $P_A(q)$ (Eq.(6)), $P_B(q)$ (Eq.(11)) and $P(q)$ (Eq.(12)), we need to obtain the limiting probability $\pi(i,j,q)$ for the Markov process $\{Y_A(t), Y_B(t), Z(t)\}$. Once $\pi(i,j,q)$ is obtained, the marginal distributions $\pi(i,q) = \sum_j \pi(i,j,q)$ and $\pi(j,q) = \sum_i \pi(i,j,q)$ can easily be computed.

We use direct numerical methods to compute $\pi(i,j,q)$. We represent the system state by (i,j,q), where i is the state of stream A, j is the state of stream B, and q is the number of packets in the system. For $i,i' \in \{1,2\}$, $j,j' \in \{3,4\}$, and $q,q' \in \{0,1,2,...,K\}$, the infinitesimal generator [6] $\mathbf{Q}$ for our system is given by

$$\mathbf{Q}_{(i,j,q)\to(i',j',q')} = \begin{cases} r_{ii'}, & \text{if } i' \neq i, j'=j, q'=q \\ r_{jj'}, & \text{if } i'=i, j'\neq j, q'=q \\ \lambda_i+\lambda_j, & \text{if } i'=i, j'=j, q'=q+1 \\ \mu, & \text{if } i'=i, j'=j, q'=q-1 \\ -r_{i\bar{i}}-r_{j\bar{j}}-\lambda_i-\lambda_j, & \text{if } i'=i, j'=j, q'=q=0 \\ -\mu-r_{i\bar{i}}-r_{j\bar{j}}, & \text{if } i'=i, j'=j, q'=q=K \\ -\mu-r_{i\bar{i}}-r_{j\bar{j}}-\lambda_i-\lambda_j, & \text{if } i'=i, j'=j, q'=q, 0<q<K \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

where r_{ij} is the transition rate of an arrival process from state i to state j, and $\bar{i}$ represents the complementary state of i. $r_{i,j}$ and $\bar{i}$ are given by $r_{12}=\alpha$, $r_{21}=\beta$, $r_{34}=\gamma$, $r_{43}=\delta$, $\bar{1}=2$, $\bar{2}=1$, $\bar{3}=4$, and $\bar{4}=3$.

Using the above infinitesimal generator, the steady state equation for our system becomes

$$\pi\mathbf{Q} = 0 \tag{17}$$

where $\pi = [\pi(1,3,0), \pi(1,4,0), \pi(2,3,0), \pi(2,4,0), \ldots, \pi(2,3,K), \pi(2,4,K)]$. By solving the above set of equations with the condition $\pi\mathbf{e} = 1$, we can easily obtain the steady-state probability $\pi(i,j,q)$.

3 Discrete-Time Case

In this section, individual packet loss probabilities are obtained for the discrete-time case. Again, we first consider multiplexing two heterogeneous streams (stream A and stream B). The extension of the discrete-time analysis for the $N(>2)$ stream case is discussed in Appendix B.

3.1 Model and Notations

The same derivation technique used earlier for continuous time is now brought to bear on the discrete-time case. Before we proceed with our analysis, we first decide the order in which arrivals and services take place and the times they occur. Without loss of generality, we assume the late arrival system with immediate access [12]. In such a system, arrivals occur just prior to the end of a time slot, and the packet in service is ejected from the service facility immediately after the beginning of a time slot (Refer to Figure 2). An arriving packet can enter the service facility if it is free, with the possibility of it being ejected almost instantaneously. Note that in this model, packet's service time is counted as the number of slot boundaries from the entering point to the service facility to the packet departure point. Therefore, even though we allow the arriving packet to be ejected almost instantaneously, its service time is counted as 1, not 0.

Consider a single first-come-first-served queue driven by two 2-state MMBP arrival processes. As in the continuous time case, the 'driving' states of stream A are labeled 1 and 2; the driving states of stream B are labeled 3 and 4 (Refer to Figure 3). Without loss of generality, it is assumed that change in the states of the arrival processes occur just prior to the end of a time slot. The sojourn times in states 1 and 2 are geometrically distributed with the mean $1/\alpha$ and $1/\beta$ slots, respectively, and the sojourn times in states 3 and 4 are also geometrically distributed with the mean $1/\gamma$ and $1/\delta$ slots, respectively. Packets arrive according to a Bernoulli process, and the probability of an arrival in a slot is p_i ($0 \leq p_i \leq 1$) in state i. Service times are geometrically distributed, and the probability of service completion in a slot, provided the server is busy, is s ($0 < s < 1$) for both stream A and stream B.

Let $n = 0, 1, 2, 3, \ldots$ denote the slot boundary numbers and t continue to denote (continuous) time. Define the step function $\lfloor t \rfloor$, where $\lfloor t \rfloor = n$, if $n \leq t < n+1$. For simplicity, we assume the slot length is equal to a unit time in the system. We observe the system just prior to the end of time slots, i.e., at $n-$ (refer to Figure 2).

As in the continuous-time case, we define $Y_A(t)$ and $Y_B(t)$ as the state of the MMBP for stream A and for stream B at time t, respectively. We also let $Z(t)$ ($0 \leq Z(t) \leq K$) denote the system state at time t. $\pi(i,q)$, $\pi(j,q)$ and $\pi(i,j,q)$ are defined as the limiting probabilities of the Markov processes $\{Y_A(n-), Z(n-)\}$, $\{Y_B(n-), Z(n-)\}$, and $\{Y_A(n-), Y_B(n-), Z(n-)\}$, respectively. Note that $i = 1, 2$, $j = 3, 4$, and $0 \leq q \leq K$.

In a discrete-time case, a packet arriving at a system whose state is K is lost. Loss of packets can also happen when simultaneous arrivals occur from streams A and B at the system state $K-1$. In such a case, we assume that a packet from stream A is lost with probability P_A, and a packet from stream B is lost with probability $P_B (= 1 - P_A)$. For a random packet discarding scheme, $P_A = P_B = 0.5$.

In the following analysis, we focus on stream A and obtain its packet loss probability. For each slot n, define the following indicator functions:

- $J(n)$: $J(n) = 1$ iff an arrival has taken place from stream B in the n^{th} slot.
- $V(n)$: $V(n) = 1$ iff in the n^{th} slot, a stream A packet is discarded when simultaneous arrivals occur from streams A and B at the system state $K-1$.

Note that, for each n, $V(n)$ is an independent Bernoulli random variable, and thus, $V(1), V(2), \ldots$ is an *iid* sequence.

Using $J(n)$ and $V(n)$, we can obtain the indicator function $U(n)$ for the state in which a stream A packet is discarded, i.e., $U(n) = 1$ iff in the n^{th} slot, stream A is in the state where a stream A packet is discarded. $U(n)$ is given by

$$U(n) = U_K(n) + U_{K-1}(n)J(n)V(n) \tag{18}$$

where $U_q(n)$ is the indicator function for the system state q, i.e., $U_q(n) = 1$ iff $Z(n-) = q$.

3.2 Analysis

First, let us derive the packet loss probability $P_{loss}(A)$ for stream A. From the definition of $P_{loss}(A)$, we have

$$\begin{aligned} P_{loss}(A) &= \lim_{t\to\infty} \frac{1}{N_A(t)} \int_0^t U(s)dN_A(s) \\ &= \lim_{t\to\infty} \frac{t}{N_A(t)} \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN_A(s). \end{aligned} \tag{19}$$

Since the term $\lim_{t\to\infty} \frac{N_A(t)}{t}$ represents the mean arrival rate of stream A, it becomes

$$\lim_{t\to\infty} \frac{N_A(t)}{t} = \frac{p_1\beta + p_2\alpha}{\alpha+\beta}. \tag{20}$$

In order to obtain the term $\lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN_A(s)$, we will use the following manner of writing the compensator for a MMBP:

$$\Lambda_A(t) = \int_0^t (I_1(s)p_1 + I_2(s)p_2)d\lfloor s \rfloor. \tag{21}$$

(The following integral and sum are the same: $\int_0^t I_1(s)d\lfloor s\rfloor = \sum_{i=0}^{\lfloor t\rfloor} I_1(i)$.) We may now apply Theorem 1 directly. (Note that 'discreteness' is accounted for via the compensator.)

$$
\begin{aligned}
\lim_{t\to\infty}\frac{1}{t}\int_0^t U(s)dN_A(s) &= \lim_{t\to\infty}\frac{1}{t}\int_0^t U(s)d\Lambda_A(s) \\
&= \lim_{t\to\infty}\frac{1}{t}\int_0^t U(s)(I_1(s)p_1 + I_2(s)p_2)d\lfloor s\rfloor \\
&= p_1\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U(n)I_1(n) + p_2\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U(n)I_2(n) \\
&= p_1\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_K(n)I_1(n) + p_1\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)V(n)I_1(n) \\
&\quad +p_2\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_K(n)I_2(n) + p_2\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)V(n)I_2(n) \\
&= p_1\pi(1,K) + p_1\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} V(n)\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)I_1(n) \\
&\quad +p_2\pi(2,K) + p_2\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} V(n)\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)I_2(n). \quad (22)
\end{aligned}
$$

In the above derivation, we used Eq.(18). For the last step, we used the fact that $\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_K(n)I_1(n)$ ($\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_K(n)I_2(n)$) represents the limiting probability for $\{Y_A(n-) = 1, Z(n-) = K\}$ ($\{Y_A(n-) = 2, Z(n-) = K\}$), and that it is equal to $\pi(1,K)$ ($\pi(2,K)$).

Note that, since we assume that a packet from stream A is discarded with probability P_A when two arrivals occur in state $K-1$, $P[V(n) = 1] = P_A$. (For a random packet discarding scheme, $P[V(n) = 1] = 0.5$.) Observe that $\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} V(n) = P_A$. Also note that $\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)I_1(n)$ represents the limiting probability that stream A is in state 1, system state is $K-1$, and an arrival from stream B occurs. It is thus equal to $p_3\pi(1,3,K-1) + p_4\pi(1,4,K-1)$. Similarly, $\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m} U_{K-1}(n)J(n)I_2(n)$ is equal to $p_3\pi(2,3,K-1) + p_4\pi(2,4,K-1)$. Then, the above equation (Eq.(22)) becomes

$$
\begin{aligned}
\lim_{t\to\infty}\frac{1}{t}\int_0^t U(s)dN_A(s) &= p_1\pi(1,K) + P_Ap_1\left(p_3\pi(1,3,K-1) + p_4\pi(1,4,K-1)\right) \\
&\quad +p_2\pi(2,K) + P_Ap_2\left(p_3\pi(2,3,K-1) + p_4\pi(2,4,K-1)\right) \\
&= p_1\pi(1,K) + p_2\pi(2,K) + P_A\{p_1p_3\pi(1,3,K-1) + p_1p_4\pi(1,4,K-1) \\
&\quad +p_2p_3\pi(2,3,K-1) + p_2p_4\pi(2,4,K-1)\} \quad (23)
\end{aligned}
$$

By substituting Eqs.(20) and (23) into Eq.(19), we have

$$
P_{loss}(A) = \frac{(\alpha+\beta)}{(p_1\beta + p_2\alpha)}[p_1\pi(1,K) + p_2\pi(2,K) + P_A\{p_1p_3\pi(1,3,K-1)
$$

$$+p_1p_4\pi(1,4,K-1)+p_2p_3\pi(2,3,K-1)+p_2p_4\pi(2,4,K-1)\}]. \qquad (24)$$

Using the same argument, we can obtain the loss probability for packets from stream B, and it is given by

$$\begin{aligned} P_{loss}(B) &= \frac{(\gamma+\delta)}{(p_3\delta+p_4\gamma)}[(p_3\pi(3,K)+p_4\pi(4,K)+P_B\{p_1p_3\pi(1,3,K-1) \\ &\quad +p_1p_4\pi(1,4,K-1)+p_2p_3\pi(2,3,K-1)+p_2p_4\pi(2,4,K-1)\}] \qquad (25) \end{aligned}$$

For the loss probability $P_{loss}(O)$ seen by an arbitrary arrival, we have

$$P_{loss}(O) = \lim_{t\to\infty} \frac{1}{N(t)} \int_0^t U(s)dN(s) = \lim_{t\to\infty} \frac{t}{N(t)} \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN(s) \qquad (26)$$

Since $N(t) = N_A(t) + N_B(t)$, we obtain

$$\lim_{t\to\infty} \frac{N(t)}{t} = \lim_{t\to\infty} \frac{N_A(t)}{t} + \lim_{t\to\infty} \frac{N_B(t)}{t} = \frac{p_1\beta+p_2\alpha}{\alpha+\beta} + \frac{p_3\delta+p_4\gamma}{\gamma+\delta} \qquad (27)$$

and

$$\begin{aligned} \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN(s) &= \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN_A(s) + \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s)dN_B(s) \\ &= p_1\pi(1,K)+p_2\pi(2,K)+p_3\pi(3,K)+p_4\pi(4,K)+p_1p_3\pi(1,3,K-1) \\ &\quad +p_1p_4\pi(1,4,K-1)+p_2p_3\pi(2,3,K-1)+p_2p_4\pi(2,4,K-1) \qquad (28) \end{aligned}$$

By substituting Eqs.(27) and (28) into Eq.(26), we can easily obtain $P_{loss}(O)$.

In order to obtain the packet loss probabilities $P_{loss}(A)$, $P_{loss}(B)$ and $P_{loss}(O)$, we need to obtain the limiting probability $\pi(i,j,q)$ of the Markov process $\{Y_A(n-), Y_B(n-), Z(n-)\}$. For this purpose, we use direct numerical methods as in the continuous-time case.

The input stream A is described by a two-state Markov chain with transition probability matrix $\mathbf{B_1}$ given by

$$\mathbf{B_1} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}. \qquad (29)$$

The transition probability matrix for the input stream B, $\mathbf{B_2}$, is given by

$$\mathbf{B_2} = \begin{pmatrix} 1-\gamma & \gamma \\ \delta & 1-\delta \end{pmatrix}. \qquad (30)$$

The aggregated input process is fully characterized by the product chain of these two independent two-state streams. The transition probability matrix $\mathbf{B}$ of the product chain is then given by the Kronecker product

$$\mathbf{B} = \mathbf{B_1} \otimes \mathbf{B_2} = \begin{pmatrix} (1-\alpha)(1-\gamma) & (1-\alpha)\gamma & \alpha(1-\gamma) & \alpha\gamma \\ (1-\alpha)\delta & (1-\alpha)(1-\delta) & \alpha\delta & \alpha(1-\delta) \\ \beta(1-\gamma) & \beta\gamma & (1-\beta)(1-\gamma) & (1-\beta)\gamma \\ \beta\delta & \beta(1-\delta) & (1-\beta)\delta & (1-\beta)(1-\delta) \end{pmatrix}. \tag{31}$$

As in the continuous-time case, we represent the system state by (i, j, q), where i is the state of stream A, j is the state of stream B, and q is the number of packets in the system. For convenience, let $\mathbf{q}$ denote the set of states for a given value of q $\{(i, j, q), i \in \{1, 2\}, j \in \{3, 4\}\}$. The transition probability matrix $\mathbf{P}$ of our system is then given in Figure 4. In Figure 4, P_{ia} represents the probability that i number of packets arrive, P_{id} represents the probability that i number of packets depart, and P_a' represents the probability that some number of packets arrive. When $Y_A = i$ $(i = 1, 2)$ and $Y_B = j$ $(j = 3, 4)$, we have $P_{0a} = (1-p_i)(1-p_j)$, $P_{1a} = p_i(1-p_j) + p_j(1-p_i)$, $P_{2a} = p_i p_j$, and $P_a' = 1 - P_{0a} = 1 - (1-p_i)(1-p_j)$. Further, $P_{0d} = 1 - s$ and $P_{1d} = s$.

By solving the set of steady state equations $\pi\mathbf{P} = \pi, \pi\mathbf{e} = 1$ where $\pi = [\pi(1,3,0), \pi(1,4,0), \pi(2,3,0), \pi(2,4,0), \ldots, \pi(2,3,K), \pi(2,4,K)]$, we can obtain the steady-state probability $\pi(i, j, q)$. Then, the marginal distribution such as $\pi(i, q) = \sum_j \pi(i, j, q)$ can easily be computed.

Note that for a discrete-time analysis, the extra complication of simultaneous arrivals is accounted for quite easily with the stochastic integral approach used in our paper. Our analysis can easily be extended to accommodate a wide range of packet discarding schemes. In fact, any state-based packet discarding schemes can be analyzed in the same fashion. As an example, a priority packet discarding scheme [22, 21], a technique frequently proposed for high-speed networks, is analyzed in the following section.

4 Priority Packet Discarding Scheme

In this section, our analysis presented in section 3 is extended to accommodate a priority packet discarding scheme. A priority packet discarding scheme is a popular congestion control technique for high-speed networks [22, 21, 18, 13]. It can be used to satisfy varying loss requirements of different classes of traffic. In general, loss-sensitive traffic such as data is given priority over loss-insensitive traffic such as voice. When network congestion occurs, varying loss requirements of different classes of traffic can be satisfied by selectively discarding low priority packets.

For voice or video traffic, a priority packet discarding scheme may be used in conjunction with an embedded coding technique [1] [9, 13, 8]. If an embedded coding technique is used for voice, the encoded information is divided into more significant bits and less significant bits. More significant bits form high priority packets, and less significant bits form low priority packets. If an embedded coding technique is used for video, low frequency components of video form high priority packets, and high frequency components of video (refinement of image) form low priority packets. With an embedded coding, packets containing more important information are given higher priority than packets containing less important information, and when network congestion occurs, packets containing less important information are discarded first.

In this section, a simple threshold-based discarding scheme [3] is considered. With this scheme, low priority packets are accepted only if the current system occupancy is less than a certain threshold.

Again it is assumed that two heterogeneous streams (stream A and stream B) are multiplexed. The same model used in section 3 is assumed. The extension of the analysis for $N(>2)$ stream case is discussed in Appendix C.

Assume that stream A has a higher priority than stream B. For the high priority stream A,

$$U(n) = U_K(n) \tag{32}$$

since arriving high priority packets are discarded only when the system size becomes K. For the low priority stream B,

$$U(n) = \sum_{q=\theta}^{K} U_q(n) \tag{33}$$

since arriving low priority packets are always discarded when the system occupancy is greater than or equal to θ. For the remaining derivation, the same analytical technique used in subsection 3.2 applies

The transition probability matrix $\mathbf{P}$ for this system is given in Figure 5. P^h_{ia} represents the probability that i number of high priority packets arrive. When $Y_A = i$ $(i = 1,2)$, we have $P^h_{0a} = (1 - p_i)$ and $P^h_{1a} = p_i$. For the rest of the notation used in Figure 5, refer to subsection 3.2.

Note that the method used in our analysis, a new stochastic integral approach, is easily applied to derive individual packet loss probabilities when a packet discarding control scheme is employed.

5 Numerical Examples

In this section, through numerical examples, the effects of traffic characteristics on the individual packet loss probabilities are investigated. It is assumed, for simplicity, that two heterogeneous

[1] Also referred to as a layered coding technique or a hierarchical coding technique.

streams (stream A and stream B) are multiplexed. We show how the burstiness of one stream affects the packet loss probabilities of each of the two multiplexed streams.

Burstiness is one of the most critical parameters in determining the network performance. A number of ways have been proposed to describe the burstiness of a traffic source (see [2] for a discussion). However, consensus is yet to be reached concerning an appropriate way to describe the burstiness of a traffic source. In keeping with our focus on Markov Modulated Arrivals, we examine the following three intuitive ways to vary the burstiness of a stream. In all three we keep the mean arrival rate of the stream *constant*. The expression for the mean arrival rate of the steam is given in Eq.(7) for the continuous-time case and in Eq.(20) for the discrete-time case.

Method 1: Keep the average sojourn times in two driving states constant, and vary the arrival rates in two states. In this case, as the difference between the arrival rates in two states increases, the burstiness of the stream also increases.

Method 2: For an Interrupted Poisson Process (IPP) stream, keep the ratio of average active period (i.e., period during which packets are generated) to average idle period (i.e., period during which no packets are generated) constant, and vary both active and idle periods. In this case, as the average active and idle period increase, the burstiness of the stream also increases.

Method 3: For an IPP stream, keep the sum of the average active and idle periods constant, and vary the average active and idle periods. Since we keep the mean arrival rate constant, a smaller active period means a greater arrival rate during an active period. In this case, as the average active period decreases (i.e., as the arrival rate during an active period increases), the burstiness of the stream increases.

Note that the first and the third methods are two ways to vary peak-to-mean ratio. The peak-to-mean ratio is the most commonly used definition of the burstiness. The second method varies the average active period. The average active period is also a widely used parameter to measure the degree of the burstiness (see, for example [7, 11]).

In the following numerical examples, in order to characterize the effects of mixing bursty streams with non-bursty streams, we introduce the concept of "self-loss" for a single stream, the packet loss incurred when a stream is multiplexed with itself. In the figures, the following notation is used to represent the loss probabilities:

- $P_{loss}(A)$: the packet loss probability of stream A when it is multiplexed with stream B,
- $P_{loss}(B)$: the packet loss probability of stream B when it is multiplexed with stream A,
- $P_{loss}(O)$: the overall packet loss probability (i.e., the loss probability for all streams),

- self-loss(A): the packet loss probability of stream A when it is multiplexed with itself, and
- self-loss(B): the packet loss probability of stream B when it is multiplexed with itself.

In the following subsections, 5.1 and 5.2, continuous and discrete-time cases are examined, respectively. Results obtained in subsections 5.1 and 5.2 are summarized in subsection 5.3.

5.1 Continuous-Time Case

Let m_A denote the mean arrival rate of stream A and m_B denote the mean arrival rate of stream B. From Eq.(7), we have

$$m_A = \frac{\lambda_1\beta + \lambda_2\alpha}{\alpha + \beta}. \tag{34}$$

Similarly, we have

$$m_B = \frac{\lambda_3\delta + \lambda_4\gamma}{\gamma + \delta}. \tag{35}$$

The offered load, ρ, is given by

$$\rho = \frac{m_A + m_B}{\mu} \tag{36}$$

Throughout the numerical examples in this subsection, we assume the maximum system size $K = 10$ (i.e., buffer size = 9), the service rate $\mu = 0.8$, and the offered load $\rho = 0.1$. We further assume that the mean arrival rates of two streams are the same. This allows us to investigate solely the effect of burstiness. $m_A = m_B = 0.04$ since the offered load $\rho = 0.1$. Note that the offered load is low in order to keep loss probabilities realistically small.

Stream A is assumed to be a Poisson stream (i.e., $\lambda_1 = \lambda_2 = 0.04$). We vary the burstiness of stream B keeping the mean arrival rate constant. Stream A is kept constant in all the figures.

In Figure 6, the mean arrival rate of stream B is kept constant and is equal to 0.04. In this figure, the burstiness of stream B is varied using the first method described above. In other words, λ_3 and λ_4 (arrival rates in two states) are varied, keeping $\frac{1}{\gamma}$ and $\frac{1}{\delta}$ (the average sojourn times in two states) constant. The bigger the difference between λ_3 and λ_4, the greater the burstiness of stream B. In this figure, the values of γ and δ are 0.01 and 0.19, respectively. The horizontal axis shows the difference between λ_4 and λ_3. At the leftmost starting point (i.e., when $\lambda_4 - \lambda_3 = 0$), stream B becomes a Poisson stream. In this figure, moving to the right, the difference between λ_4 and λ_3 becomes larger, and thus, the burstiness of stream B increases.

Several observations can be made from Figure 6. At the leftmost starting point, the packet loss probabilities for stream A and stream B are the same since, at this point, stream B is also a Poisson stream (i.e., $\lambda_3 = \lambda_4 = 0.04$). As the burstiness of stream B increases, both loss probability of stream B and loss probability of stream A increase. From this, it can be concluded that an increase in the burstiness of one stream negatively affects the stream itself and also the other multiplexed stream.

Next, compare the $P_{loss}(A)$ curve with the self-loss(A) curve. The self-loss(A) (i.e., the loss probability of stream A when it is multiplexed with an identical stream) is always smaller than the $P_{loss}(A)$ (i.e., the loss probability of stream A when it is multiplexed with stream B). In other words, a Poisson stream (stream A) is penalized by sharing a buffer with a bursty stream (stream B), as opposed to sharing a buffer with another Poisson stream. This is because bursty stream causes buffer buildups, blocking the Poisson stream.

Compare the $P_{loss}(B)$ curve with the self-loss(B) curve. The $P_{loss}(B)$ is always smaller than the self-loss(B). This shows that a bursty stream (stream B) gains (i.e., $P_{loss}(B) <$ self-loss(B)) by sharing a buffer with a Poisson stream (stream A), as opposed to sharing a buffer with another bursty stream. This is because a Poisson stream does not cause as much buffer buildup as a bursty stream does, and thus, a stream loses less packets when it is multiplexed with a Poisson stream than when it is multiplexed with a bursty stream. From this, it can be concluded that the traffic mix has a significant effect on the packet loss probabilities.

Figure 6 also shows that when two different traffic streams are multiplexed, the stream with the smaller self-loss probability is penalized. In this case, self-loss(A) is smaller than self-loss(B), and stream A is penalized (i.e., $P_{loss}(A) >$ self-loss(A)). Furthermore, the bigger the difference between two self-loss probabilities, the greater the penalty.

In Figure 7, the mean arrival rate of stream B is kept constant and is equal to 0.04. In this figure, an IPP is used for stream B (i.e., $\lambda_3 = 0$). λ_4 is equal to 0.8. The burstiness of stream B is varied using the second method described above. In other words, we vary $\frac{1}{\gamma}$ (the average idle period of stream B) and $\frac{1}{\delta}$ (the average active period of stream B), keeping their ratio $\frac{\delta}{\gamma}$ constant. In this case, the longer the average active (or idle) length, the greater the burstiness. In this figure, the value of $\frac{\delta}{\gamma}$ is equal to 19. The horizontal axis shows the average idle period. In this figure, moving to the right, both average active period and idle period increase, and thus, the burstiness of stream B increases. Similar results to those in Figure 6 are found in this figure. An increase in the burstiness of one stream negatively affects both the stream itself and the other stream multiplexed. When a bursty stream and a Poisson stream are multiplexed together, the Poisson stream is penalized, and the bursty stream benefits. The stream with the smaller self-loss probability is penalized, and the bigger the difference between two self-loss probabilities, the greater the penalty.

In Figure 8, the mean arrival rate of stream B is kept constant and is equal to 0.04. In this figure, an IPP is used for stream B (i.e., $\lambda_3 = 0$). The burstiness of stream B is varied using the third method described above. In other words, we vary $\frac{1}{\delta}$ (the average active period) and $\frac{1}{\gamma}$ (the average idle period), keeping $\frac{1}{\gamma} + \frac{1}{\delta}$ (the sum of average active and idle period) constant. In this figure, $\frac{1}{\gamma} + \frac{1}{\delta} = 100$. λ_4 can be computed using Eq.(35). For instance, when the idle period ($\frac{1}{\gamma}$) is 20, the active period ($\frac{1}{\delta}$) is 80 and λ_4 is 0.05, and when the idle period is 80, the active period is 20 and λ_4 is 0.2. In this figure, moving to the right, the active period ($= 1/\delta$) decreases and λ_4 increases, therefore, the burstiness of stream B increases. Again, similar observations made for Figures 6 and 7 can be made for this figure.

5.2 Discrete-Time Case

From Eq.(20), we have

$$m_A = \frac{p_1\beta + p_2\alpha}{\alpha + \beta}. \tag{37}$$

Similarly, we have

$$m_B = \frac{p_3\delta + p_4\gamma}{\gamma + \delta}. \tag{38}$$

The offered load, ρ, is given by

$$\rho = \frac{m_A + m_B}{s} \tag{39}$$

As in the continuous-time case, we assume $K = 10$, $s = 0.8$, $\rho = 0.1$, and $m_A = m_B = 0.04$. Stream A is assumed to follow a geometric arrival process (i.e., $p_1 = p_2$), and we vary the burstiness of stream B keeping its mean arrival rate constant. Stream A is kept constant in all the figures. A random packet discarding scheme is used.

Numerical results (Figures 9 through 11) for the discrete-time case are very similar to those for the continuous-time case. In Figure 9, the burstiness of stream B is varied using the first method described earlier; arrival rates in two driving states, p_3 and p_4, are varied keeping the average idle period ($\frac{1}{\gamma}$) and the average active period ($\frac{1}{\delta}$) constant. As in the continuous-time case, the values of γ and δ are 0.01 and 0.19, respectively. Figure 9 shows results very similar to Figure 6 (continuous-time case).

In Figure 10, a discrete-time version of IPP is used for stream B (i.e., $p_3 = 0$). p_4 is equal to 0.8. The burstiness of stream B is varied using the second method described earlier; the average idle period ($\frac{1}{\gamma}$) and the average active period ($\frac{1}{\delta}$) are varied keeping their ratio $\frac{\delta}{\gamma}$ constant. As in the continuous-time case, we let $\frac{\delta}{\gamma} = 19$. The behavior shown here is similar to that seen in Figure 7.

In Figure 11, a discrete-time version of IPP is used for stream B (i.e., $p_3 = 0$). The burstiness of stream B is varied using the third method; $\frac{1}{\gamma}$ and $\frac{1}{\delta}$ are varied keeping $\frac{1}{\gamma} + \frac{1}{\delta}$ constant. As in the continuous-time case, we let $\frac{1}{\gamma} + \frac{1}{\delta} = 100$. p_4 can be computed using Eq.(38). Again, Figure 11 shows very similar results to Figure 8 (continuous-time case).

5.3 Summary

In the numerical results section, the effects of traffic characteristics and the traffic mix on the packet loss probability of each of the input streams were investigated. For both continuous and discrete-time cases, the following observations were made:

- An increase in the burstiness of one stream results in an increase in the packet loss probabilities of that stream and of others which are multiplexed together.
- When two different traffic streams are multiplexed, the less bursty stream is always penalized, and the more bursty stream always benefits.
- When two different traffic streams are multiplexed together, the stream with the smaller self-loss probability is penalized. The bigger the difference between two self-loss probabilities, the greater the penalty.

Finally, note that the differences between individual loss probabilities are significant in all the figures presented in numerical example section (Figures 6 through 11). In our numerical examples, we assumed that the mean arrival rates of streams A and B are the same. However, the difference between the packet loss probabilities for stream A and B is often an order of magnitude or greater. This shows that the overall packet loss probability may not provide sufficient insight when heterogeneous traffic sources are multiplexed.

The importance of individual loss probabilities is better explained through an example. Consider admission control [7, 11, 20]. Admission control decides whether to accept or reject a new call based on whether the required performance can be maintained. If the overall packet loss probability is used as a criterion in admission control when heterogeneous traffic sources are multiplexed, the GOS of the new call may not be guaranteed. This is because, depending on the burstiness of a new coming call, its packet probability may be significantly larger than the overall packet loss probability. For example, in Figure 6, the difference between the packet loss probability of stream A and the overall packet loss probability is about an order of magnitude when the difference between the burstiness of two streams are the largest.

6 Conclusion and Future Work

In this paper, we considered a queueing system with a finite buffer and several heterogeneous arrival streams to investigate how heterogeneity affects packet loss in individual streams. We examined the class of MMA streams both in continuous time (a MMPP) and in discrete time (a MMBP) and presented an exact analysis of individual packet loss for MMA streams. We also derived individual packet loss probabilities for a well-known packet discarding scheme. Our method of analysis used a new stochastic integral approach, allowing both continuous and discrete time as well as the packet discarding scheme to be treated similarly.

Our emphasis was on examining the effect of burstiness of traffic streams on packet loss. The concept of self-loss for a single stream, the packet loss incurred when a stream is multiplexed with itself, was introduced to study the effects of multiplexing heterogeneous arrival streams on individual packet loss. The results showed that an increase in the burstiness of one stream negatively affects both the stream itself and the other stream which is multiplexed together. It was also shown that when two different traffic streams are multiplexed together, the less bursty stream is always penalized, and the more bursty stream always benefits. Furthermore, the stream with the smaller self-loss probability is penalized, and the bigger the difference between two self-loss probabilities, the greater the penalty.

The present model analyzes one node and obtains individual packet loss probabilities. Possible future work would consider a network-of-queues and obtain individual packet loss probabilities on the end-to-end basis. Also, several implementation issues are raised when it is necessary to perform the GOS calculation in real-time; if the streams are MMA it would be interesting to develop efficient computation techniques.

Appendix A: Individual Loss Probabilities for Multiple Stream Inputs Continuous-Time Case

In this appendix, we extend the continuous-time analysis presented in section 2 to the case in which more than two heterogeneous input streams are multiplexed. We obtain the packet loss probability for each input stream. The same analytical techniques used in Section 2 can be applied for a multiple stream input case. In the following, we denote the number of input streams as N, and the maximum buffer size as $K-1$ (i.e., the maximum system size of K).

The arrival process of each input stream is assumed to be a 2-state MMPP. The "driving" states of each MMPP are labeled 1 and 2. For the stream i, the transition rate from state 1 to state 2 is denoted by α_i, and the transition rate from state 2 to state 1 is denoted by β_i. The packet arrivals when the stream i is in state 1 follow a Poisson process with rate $\lambda_{1,i}$, whereas the packet arrival rate in state 2 is $\lambda_{2,i}$. The service times of packets are exponentially distributed with rate μ.

Let $N_i(t)$ denote the cumulative number of arrivals from the i^{th} stream in the time interval $[0,t]$, and $U_K(t)$ denote the indicator function for the system state K at time $t-$ (i.e., $U_K(t)=1$, iff $Z(t-)=K$). Let $P_{loss}(i)$ denote the long-term loss probability for the i^{th} stream. Then, from Eq.(6), we have

$$P_{loss}(i) = \lim_{t\to\infty} \frac{t}{N_i(t)} \lim_{t\to\infty} \frac{1}{t} \int_0^t U_K(s)dN_i(s). \tag{40}$$

Also, from Eq.(7), we have

$$\lim_{t\to\infty} \frac{N_i(t)}{t} = \frac{\lambda_{1,i}\beta_i + \lambda_{2,i}\alpha_i}{\alpha_i+\beta_i}, \tag{41}$$

and from Eq.(9), we obtain

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t U_K(s)dN_i(s) = \lambda_{1,i}\pi(1_i,K) + \lambda_{2,i}\pi(2_i,K) \tag{42}$$

Here, $\pi(1_i,K)$ and $\pi(2_i,K)$ are given by the following:

$$\pi(1_i,K) = \sum_{j_1=1}^{2}\sum_{j_2=1}^{2}\cdots\sum_{j_N=1}^{2} \pi(j_1,j_2,\cdots,j_{i-1},1,j_{i+1},\cdots,j_N,K)$$

$$\pi(2_i,K) = \sum_{j_1=1}^{2}\sum_{j_2=1}^{2}\cdots\sum_{j_N=1}^{2} \pi(j_1,j_2,\cdots,j_{i-1},2,j_{i+1},\cdots,j_N,K)$$

By substituting Eqs.(41) and (42) into Eq.(40), we can easily obtain the packet loss probability for stream i, $P_{loss}(i)$. For the overall packet loss probability, $P_{loss}(O)$, from Eq.(12), we have

$$P_{loss}(O) = \lim_{t\to\infty} \frac{t}{N(t)} \lim_{t\to\infty} \frac{1}{t} \int_0^t U_K(s)dN(s) \tag{43}$$

Since $N(t) = \sum_{i=1}^{N} N_i(t)$, we have

$$\lim_{t\to\infty} \frac{N(t)}{t} = \sum_{i=1}^{N} \lim_{t\to\infty} \frac{N_i(t)}{t} = \sum_{i=1}^{N} \frac{\lambda_{1,i}\beta_i + \lambda_{2,i}\alpha_i}{\alpha_i + \beta_i}, \tag{44}$$

and

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t U_K(s)dN(s) = \sum_{i=1}^{N} \lim_{t\to\infty} \frac{1}{t}\int_0^t U_K(s)dN_i(s) = \sum_{i=1}^{N}\{\lambda_{1,i}\pi(1_i,K) + \lambda_{2,i}\pi(2_i,K)\}. \tag{45}$$

By substituting Eqs.(44) and (45) into Eq.(43), we can easily obtain the overall loss probability.

To solve for the limiting probability $\pi(j_1, j_2, \cdots, j_N, q)$, we again use a direct numerical solution. We represent the system state by $(j_1, j_2, \cdots, j_N, q)$, where j_i ($j_i = 1,2$) is the state of the i^{th} stream ($i = 1, 2, \cdots, N$), and q is the number of packets in the system. Then, for $i, k \in \{1, 2, \cdots, N\}, j_i, j_i' \in \{1,2\}$, $q, q' \in \{0, 1, 2, ..., K\}$, the infinitesimal generator $\mathbf{Q}$, is given by

$$\mathbf{Q}_{(j_1,j_2,\cdots,j_N,q)\to(j_1',j_2',\cdots,j_N',q')} = \begin{cases} r_{j_ij_i'}, & \text{if } j_i' \neq j_i, j_k' = j_k \text{ for all } k \neq i, q' = q \\ \sum_{i=1}^{N} \lambda_{j_i,i}, & \text{if } j_i' = j_i \text{ for all } i, q' = q+1 \\ \mu, & \text{if } j_i' = j_i \text{ for all } i, q' = q-1 \\ -\sum_{i=1}^{N} r_{j_i\bar{j}_i} - \sum_{i=1}^{N} \lambda_{j_i,i}, & \text{if } j_i' = j_i \text{ for all } i, q' = q = 0 \\ -\mu - \sum_{i=1}^{N} r_{j_i\bar{j}_i}, & \text{if } j_i' = j_i \text{ for all } i, q' = q = K \\ -\mu - \sum_{i=1}^{N} r_{j_i\bar{j}_i} - \sum_{i=1}^{N} \lambda_{j_i,i}, & \text{if } j_i' = j_i \text{ for all } i, q' = q, 0 < q < K \\ 0, & \text{otherwise} \end{cases} \tag{46}$$

where $r_{j_ij_i'}$ is the transition rate of an arrival process from state j_i to state j_i', and $\bar{i}$ represents the complementary state of i (i.e., $\bar{1} = 2, \bar{2} = 1$). $r_{j_ij_i'}$ is given by $r_{j_ij_i'} = \begin{cases} \alpha_i, & \text{if } j_i = 1, j_i' = 2 \\ \beta_i, & \text{if } j_i = 2, j_i' = 1 \end{cases}$.

Appendix B: Individual Loss Probabilities for Multiple Stream Inputs Discrete-Time Case

We now extend the discrete-time analysis presented in sections 3 to the case in which more than two heterogeneous input streams are multiplexed. We obtain the packet loss probability for each input stream. In the following, we denote the number of input streams as N, and the maximum buffer size as $K-1$ (i.e., the maximum system size of K).

The arrival process of each input stream is assumed to be a 2-state MMBP. The "driving" states of each MMBP are labeled 1 and 2. For the stream i, transition rates between states are α_i (from state 1 to state 2) and β_i (from state 2 to state 1). Packets arrive according to a Bernoulli process, and for the stream i, the probability of having an arrival in a slot is $p_{1,i}$ when it is in state 1 and $p_{2,i}$ when it is in state 2. Service times are geometrically distributed, and the probability of service completion in a slot, provided the server is busy, is s.

We focus on the i^{th} stream, and we define the following indicator functions for each slot n to obtain the packet loss probability for the i^{th} stream:

- $J_l(n)$: $J_l(n) = 1 \ iff \ l$ number of arrivals have taken place from the remaining $N-1$ streams in the n^{th} slot. Note that, when $J_l(n) = 1$, a total of $l+1$ arrivals (including the arrival from the i^{th} stream) have taken place in the n^{th} slot.
- $V_{l,q}(n)$: $V_{l,q}(n) = 1 \ iff$ in the n^{th}slot, a packet from the i^{th} stream is discarded when a total of l packets arrive at the system state q.

Let $U(n)$ denote the indicator function for the state in which a packet from the i^{th} stream is discarded in the n^{th} slot. Then, we have

$$\begin{aligned} U(n) &= U_K(n) \\ &+U_{K-1}(n)\{J_1(n)V_{2,K-1}(n) + J_2(n)V_{3,K-1}(n) + \cdots + J_{N-1}(n)V_{N,K-1}(n)\} \\ &+U_{K-2}(n)\{J_2(n)V_{3,K-2}(n) + J_3(n)V_{4,K-2}(n) + \cdots + J_{N-1}(n)V_{N,K-2}(n)\} \\ &\vdots \\ &+U_{K-N+1}(n)J_{N-1}(n)V_{N,K-N+1}(n). \end{aligned} \tag{47}$$

For the random packet discarding scheme, we have

$$P[V_{l,q}(n) = 1] = \lim_{m\to\infty} \frac{1}{m}\sum_{n=0}^{m} V_{l,q}(n) = \frac{l-(K-q)}{l} \tag{48}$$

This is because, out of l packets, only $K-q$ number of packets are accepted and the remaining $l-(K-q)$ number of packets are lost. For the remaining derivation, the same analytical technique used in section 3.2 can apply.

As in section 3.2, a direct numerical method can be used to solve for the limiting probabilities. The transition probability matrix $\mathbf{P}$ is given in Figure 12. In Figure 12, the matrix $\mathbf{B}$ is given by the Kronecker product

$$\mathbf{B} = \mathbf{B_1} \otimes \mathbf{B_2} \otimes \cdots \otimes \mathbf{B_N} \tag{49}$$

where

$$\mathbf{B_i} = \begin{pmatrix} 1-\alpha_i & \alpha_i \\ \beta_i & 1-\beta_i \end{pmatrix}. \tag{50}$$

$P_{\{\geq ia\}}$ in Figure 12 represents the probability that at least i number of packets arrive. For the rest of the notation in Figure 12, same definition used in subsection 3.2 is used.

Appendix C: Individual Loss Probabilities for Multiple Stream Inputs Priority Packet Discarding Scheme

Now, the analysis of a priority packet discarding scheme for two arrival streams is extended to more than two heterogeneous input streams: N heterogeneous sources are multiplexed into a single first-come-first-served queue whose maximum capacity is $K-1$ (i.e., the maximum system size is K). For the arrival process of each stream, refer to Appendix B. Again, a threshold-based priority discarding scheme is considered. It is assumed that if the current system occupancy is less than the threshold, no low priority packets are discarded even if accepting all the low priority packets exceeds the threshold, as opposed to only accepting low priority packets up to the threshold.

Let q denote the system occupancy and θ denote the threshold. Assume that among N streams, H number of streams are high priority streams, and the rest of the streams are low priority streams. For the i^{th} stream, again, let $U(n)$ denote the indicator function for the state in which the stream i packet is discarded in the n^{th} slot, and define the following indicator functions for each slot n:

- $J_l(n)$: $J_l(n) = 1$ iff l number of arrivals have taken place from the remaining $N-1$ streams in the n^{th} slot. Note that, when $J_l(n) = 1$, a total of $l+1$ (including the arrival from the i^{th} stream) number of arrivals have taken place in the n^{th} slot.
- $J_l^h(n)$: $J_l^h(n) = 1$ iff l number of arrivals have taken place from the remaining $H-1$ high priority streams. Note that, when $J_l(n) = 1$, a total of $l+1$ (including the arrival from the i^{th} stream) number of arrivals have taken place from the high priority streams in the n^{th} slot.
- $V_{l,q}(n)$: $V_{l,q}(n) = 1$ iff in the n^{th} slot, the i^{th} stream packet is discarded when a total of l number of arrivals occur at the system state q.

First, let us consider the case when $\theta \leq K - N + 1$. Here, no packet will be discarded when $q < \theta$. When the stream i is a low priority stream,

$$U(n) = \sum_{q=\theta}^{K} U_q(n) \tag{51}$$

since arriving low priority packets are always discarded when the queue occupancy is greater than or equal to θ. When the stream i is a high priority stream,

$$\begin{aligned} U(n) &= U_K(n) \\ &\quad + U_{K-1}(n)\{J_1^h(n)V_{2,K-1}(n) + J_2^h(n)V_{3,K-1}(n) + \cdots + J_{H-1}^h(n)V_{H,K-1}(n)\} \\ &\quad + U_{K-2}(n)\{J_2^h(n)V_{3,K-2}(n) + J_3^h(n)V_{4,K-2}(n) + \cdots + J_{H-1}^h(n)V_{H,K-2}(n)\} \\ &\quad \vdots \\ &\quad + U_{K-H+1}(n)J_{H-1}^h(n)V_{H,K-H+1}(n). \end{aligned} \tag{52}$$

When $q = K$, arriving high priority packets are always discarded, and when $K - H + 1 \leq q \leq K - 1$, arriving high priority packets are discarded only if the number of arrivals from high priority streams exceeds the available buffer space. We assume that when the number of arrivals from high priority streams exceeds the available buffer space, the packets to be discarded are randomly picked. Therefore,

$$P[V_{l,q}(n) = 1] = \lim_{m \to \infty} \frac{1}{m} \sum_{n=0}^{m} V_{l,q}(n) = \frac{l - (K - q)}{l} \tag{53}$$

since out of l packets, only $K - q$ number of packets will be accepted and the remaining $l - (K - q)$ number of packets will be discarded.

Now, let us consider the case when $\theta > K - N + 1$. In this case, when $K - N + 1 \leq q < \theta$, if the number of arrivals exceeds the available buffer space, packets will be discarded randomly. Therefore, when the stream i is a low priority stream,

$$\begin{aligned} U(n) &= \sum_{q=\theta}^{K} U_q(n) \\ &\quad + U_{\theta-1}(n)\{J_{K-\theta+1}(n)V_{K-\theta+2,\theta-1}(n) + J_{K-\theta+2}(n)V_{K-\theta+3,\theta-1}(n) + \cdots + J_{N-1}(n)V_{N,\theta-1}(n)\} \\ &\quad + U_{\theta-2}(n)\{J_{K-\theta+2}(n)V_{K-\theta+3,\theta-2}(n) + J_{K-\theta+3}(n)V_{K-\theta+4,\theta-2}(n) + \cdots + J_{N-1}(n)V_{N,\theta-2}(n)\} \\ &\quad \vdots \\ &\quad + U_{K-N+1}(n)J_{N-1}(n)V_{N,K-N+1}(n). \end{aligned} \tag{54}$$

When $q \geq \theta$, arriving low priority packets are always discarded (the first line on the right hand side), and when $K - N + 1 \leq q < \theta$, arriving packets are discarded randomly when the number of arrivals exceeds the available buffer space (the rest of the terms on the right hand side).

When the stream i is a high priority stream, we need to consider the following two cases: (1) $\theta \leq K - H + 1$ and (2) $\theta > K - H + 1$. For the case when $\theta \leq K - H + 1$,

$$\begin{aligned} U(n) &= U_K(n) \\ &\quad + U_{K-1}(n)\{J_1^h(n)V_{2,K-1}(n) + J_2^h(n)V_{3,K-1}(n) + \cdots + J_{H-1}^h(n)V_{H,K-1}(n)\} \\ &\quad + U_{K-2}(n)\{J_2^h(n)V_{3,K-2}(n) + J_3^h(n)V_{4,K-2}(n) + \cdots + J_{H-1}^h(n)V_{H,K-2}(n)\} \\ &\quad \vdots \\ &\quad + U_{K-H+1}(n)J_{H-1}^h(n)V_{H,K-H+1}(n) \\ &\quad + U_{\theta-1}(n)\{J_{K-\theta+1}(n)V_{K-\theta+2,\theta-1}(n) + J_{K-\theta+2}(n)V_{K-\theta+3,\theta-1}(n) + \cdots + J_{N-1}(n)V_{N,\theta-1}(n)\} \\ &\quad + U_{\theta-2}(n)\{J_{K-\theta+2}(n)V_{K-\theta+3,\theta-2}(n) + J_{K-\theta+3}(n)V_{K-\theta+4,\theta-2}(n) + \cdots + J_{N-1}(n)V_{N,\theta-2}(n)\} \\ &\quad \vdots \\ &\quad + U_{K-N+1}(n)J_{N-1}(n)V_{N,K-N+1}(n). \end{aligned} \tag{55}$$

When $q = K$, arriving high priority packets are always discarded; when $K - H + 1 \leq q \leq K - 1$, arriving high priority packets are discarded only if the number of arrivals from high priority streams exceeds the available buffer space; and when $K - N + 1 \leq q < \theta$, arriving packets are discarded randomly when the number of arrivals exceeds the available buffer space.

For the case when $\theta > K - H + 1$,

$$
\begin{aligned}
U(n) \quad = \quad & U_K(n) \\
& + U_{K-1}(n)\{J_1^h(n)V_{2,K-1}(n) + J_2^h(n)V_{3,K-1}(n) + \cdots + J_{H-1}^h(n)V_{H,K-1}(n)\} \\
& + U_{K-2}(n)\{J_2^h(n)V_{3,K-2}(n) + J_3^h(n)V_{4,K-2}(n) + \cdots + J_{H-1}^h(n)V_{H,K-2}(n)\} \\
& \vdots \\
& + U_\theta(n)\{J_{K-\theta}^h(n)V_{K-\theta+1,\theta}(n) + J_{K-\theta+1}^h(n)V_{K-\theta+2,\theta}(n) + \cdots + J_{H-1}^h(n)V_{H,\theta}(n)\} \\
& + U_{\theta-1}(n)\{J_{K-\theta+1}(n)V_{K-\theta+2,\theta-1}(n) + J_{K-\theta+2}(n)V_{K-\theta+3,\theta-1}(n) + \cdots + J_{N-1}(n)V_{N,\theta-1}(n)\} \\
& + U_{\theta-2}(n)\{J_{K-\theta+2}(n)V_{K-\theta+3,\theta-2}(n) + J_{K-\theta+3}(n)V_{K-\theta+4,\theta-2}(n) + \cdots + J_{N-1}(n)V_{N,\theta-2}(n)\} \\
& \vdots \\
& + U_{K-N+1}(n)J_{N-1}(n)V_{N,K-N+1}(n). \qquad (56)
\end{aligned}
$$

When $q = K$, arriving high priority packets are always discarded; when $\theta \leq q \leq K - 1$, arriving high priority packets are discarded only if the number of arrivals from high priority streams exceeds the available buffer space; and when $K - N + 1 \leq q < \theta$, arriving packets are discarded randomly when the number of arrivals exceeds the available buffer space. $P[V_{l,q}(n) = 1]$ is given in Eq.(53). For the remaining derivation, the same analytical technique used in subsection 3.2 can apply.

References

[1] D. Anick, D. Mitra and M. M. Sondhi, "Stochastic Theory of a Data-handling System with Multiple Source," Bell Syst. Tech. J., pp.1871-1894, October 1982.

[2] J. J. Bae and T. Suda, "Survey of Traffic Control Schemes and Protocols in ATM Networks," Proceedings of the IEEE, Vol.79, No.2, pp.170-189, February 1991.

[3] K. Bala, Israel Cidon and K. Shoraby, "Congestion Control for High Speed Packet Switched Network," IEEE INFOCOM'90, pp.520-526.

[4] P. Bremaud, Point Processes and Queues: Martingale Dynamics, Springer-Verlag, New York, 1981.

[5] I.Cidon, I.Gopal, M.Gopal and S.Kutten, "Distributed Control in PARIS", *Proc. 9th Annual Symposium on Principles of Distributed Computing*, Quebec, Canada, 1990.

[6] E.Cinlar, *Introduction to Stochastic Processes*, Prentice Hall, 1975.

[7] G. Gallassi, G. Rigolio and L. Fratta, "ATM: Bandwidth Assignment and Bandwidth Enforcement Policies," Proc. IEEE GLOBECOM'89, pp.49.6.1-49.6.6.

[8] M. Ghanbari, "Two-Layer Coding of Video Signals for VBR Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.771-781. June 1989.

[9] D. J. Goodman, "Embedded DPCM for variable bit rate transmission," IEEE Trans. Commun., Vol.COM-28, pp.1040-1046, July 1980.

[10] H. Heffes and D. Lucantoni,"A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," IEEE J. Select. Areas Commun., Vol.SAC-4, No.6, pp.856-868, September 1986.

[11] M. Hirano and N. Watanabe, "Characteristics of a Cell Multiplexer for Bursty ATM Traffic," Proc. IEEE ICC'89, pp.13.2.1-13.2.5.

[12] J. J. Hunter, Mathematical Techniques of Applied Probability, Vol.II, Academic Press, New York, pp.193, 1983. -

[13] F. Kishino, K. Manabe, Y. Hayashi and H. Yasuda, "Variable Bit-Rate Coding of Video Signals for ATM Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.801-806, June 1989.

[14] R. S. Liptser and A. N. Shiryaev, Statistics of Random Processes, Vols.I,II, Springer-Verlag, New York, 1978.

[15] K. S. Meier-Hellstern, "The Analysis of a Queue Arising in Overflow Models," IEEE Trans. Commun., Vol.37, No.4, pp.367-372, April 1989.

[16] R. Nagarajan, J.F. Kurose and D. Towsley, "Approximation Techniques for Computing Packet Loss in Finite-Buffered Voice Multiplexers", *IEEE J. Selected Areas in Communications*, Vol. 9, No. 3, pp. 368-377, 1991.

[17] M. F. Neuts, Structured Stochastic Matrices of M/G/1 Type And Their Applications, Marcel Dekker, Inc., New York, 1989.

[18] D. W. Petr, L. A. DaSilva, Jr., and V. S. Frost, "Priority discarding of Speech in Integrated Packet Networks," IEEE J. Select. Areas Commun., Vol.7, No.5, pp.644-659, June 1989.

[19] W. A. Rosenkrantz and R. Simha, "Some Theorems on Conditional Pasta: A Stochastic Integral Approach", *Operations Research Letters*, to appear.

[20] G. M. Woodruff, R. G. Rogers, and P. S. Richards, "A Congestion Control Framework for High-Speed Integrated Packetized Transport," Proc. IEEE GLOBECOM'88, pp.7.1.1-7.1.5.

[21] G. M. Woodruff and R. Kositpaiboon, "Multimedia Traffic Management Principles for Guaranteed ATM Network Performance," IEEE J. Select. Areas Commun., Vol.8, No.3, pp.437-446, April 1990.

[22] N. Yin, S-Q. Li and T. E. Stern, "Data Performance in an Integrated Packet Voice/Data System Using Voice Congestion Control," Proc. IEEE GLOBECOM'88, pp.16.4.1-16.4.5.

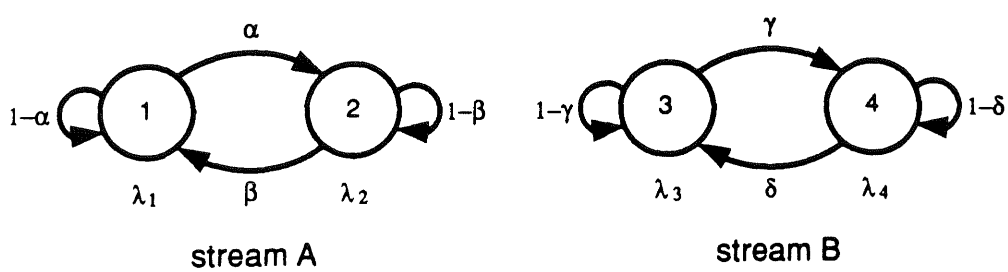

Figure 1. 2-State MMPP

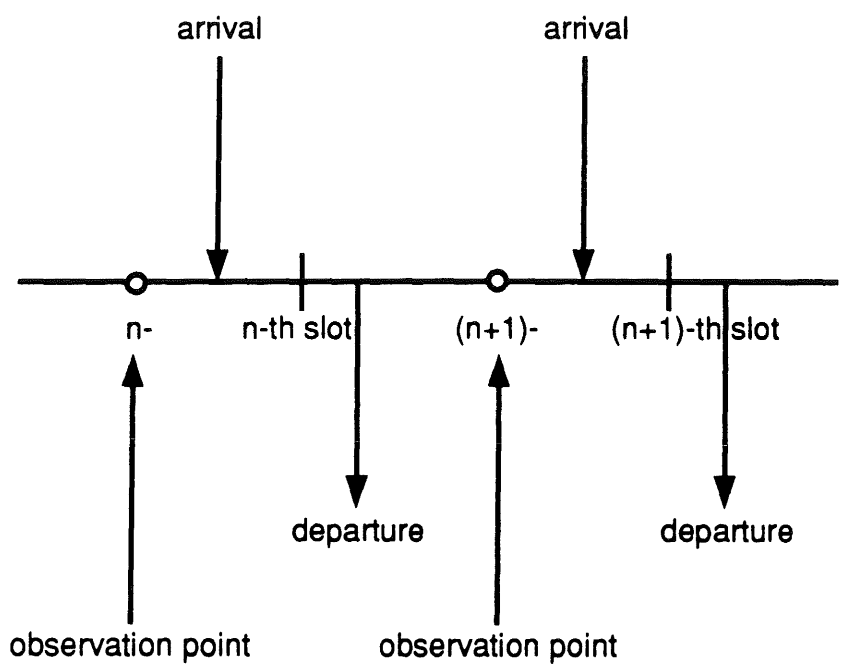

Figure 2. Late Arrival System

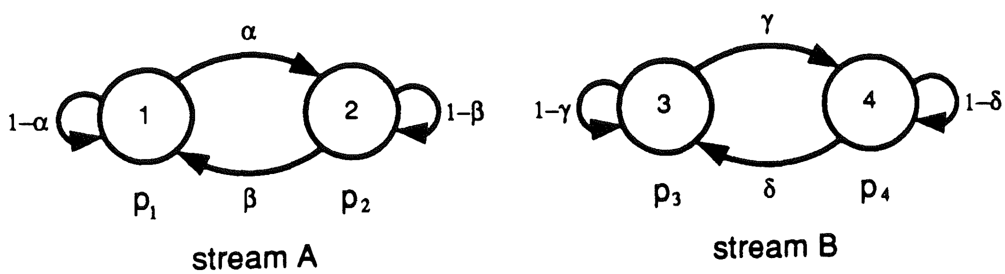

Figure 3. 2-State MMBP

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} 0 & 1 & 2 & 3 & \cdots & K\text{-}2 & K\text{-}1 & K \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ K\text{-}1 \\ K \end{array} &
\begin{pmatrix}
\{P_{1a}P_{1d}+P_{0a}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & P_{2a}P_{0d}\,B & & & & & \\
P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & P_{2a}P_{0d}\,B & & & & \\
 & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & P_{2a}P_{0d}\,B & & & \\
 & & \ddots & \ddots & \ddots & \ddots & & \\
 & & & & \vdots & & & \\
 & & & & & & & \\
 & & & & & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P'_{a}P_{1d}\}B & P'_{a}P_{0d}B \\
 & & & & & & P_{1d}B & P_{0d}B
\end{pmatrix}
\end{array}
$$

Figure 4. Transition Probability Matrix for 2-stream case (Discrete Time)

$$
\begin{array}{c|ccccccccccc}
 & 0 & 1 & 2 & \cdots & \theta-1 & \theta & \theta+1 & \theta+2 & \cdots & K-1 & K \\
\hline
0 & \{P_{1a}P_{1d}+P_{0a}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & P_{2a}P_{0d}B & & & & & & & & \\
1 & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & P_{2a}P_{0d}B & & & & & & & \\
2 & & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & \ddots & \ddots & & & & & & \\
\cdot & & & & \vdots & & & & & & & \\
\cdot & & & & & & & & & & & \\
\cdot & & & & & & & & & & & \\
\theta & & & & & P^{h}_{0a}P_{1d}B & \{P^{h}_{0a}P_{0d}+P^{h}_{1a}P_{1d}\}B & P^{h}_{1a}P_{0d}B & \ddots & & & \\
\cdot & & & & & & \ddots & \ddots & & & & \\
\cdot & & & & \vdots & & & & & & & \\
\cdot & & & & & & & & & & & \\
K & & & & & & & & & & P_{1d}B & P_{0d}B
\end{array}
$$

Figure 5. Transition Probability Matrix P for the Sytem with Threshold-Based Priority Packet Discarding (2-stream, Discrete Time)

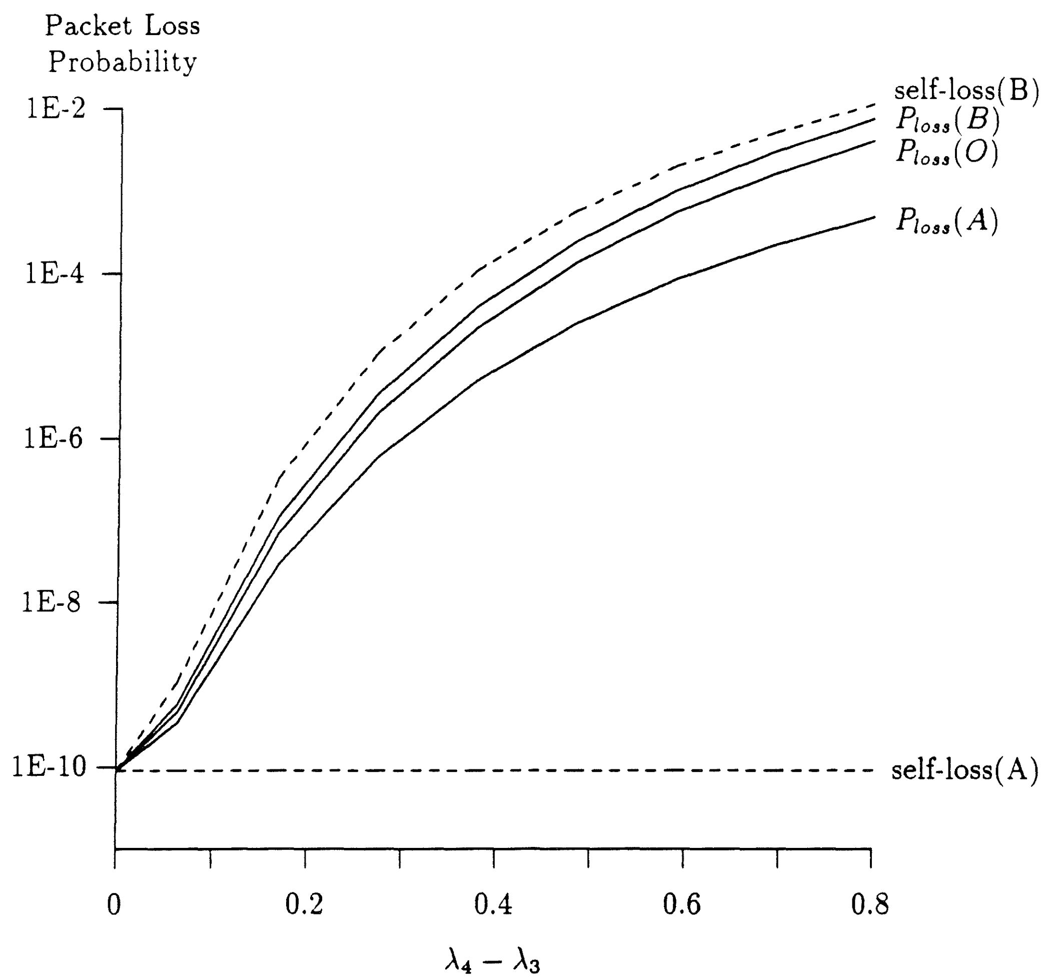

Figure 6. Effect of Burstiness on Packet Loss Probabilities
Continuous Time, The First Method
($\lambda_1 = \lambda_2 = 0.04$, $\gamma = 0.01$, $\delta = 0.19$)

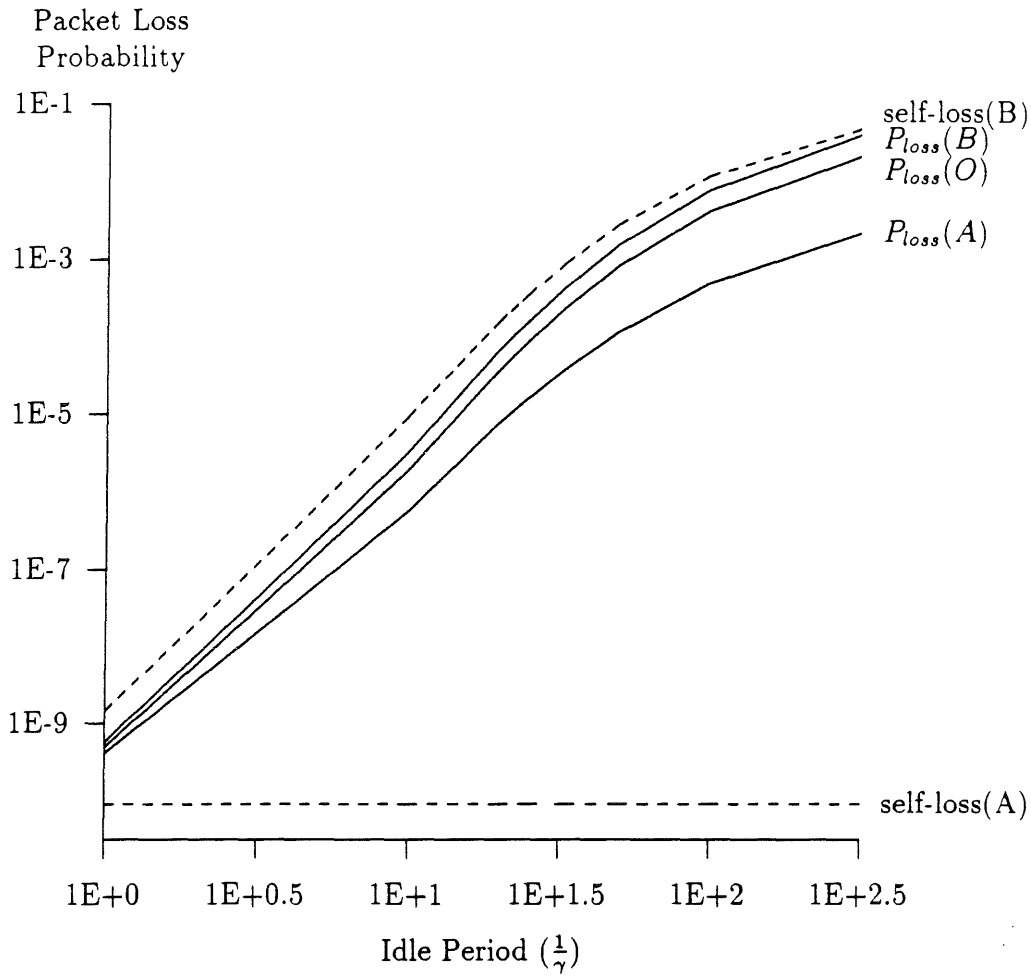

Figure 7. Effect of Burstiness on Packet Loss Probabilities
Continuous Time, The Second Method
($\lambda_1 = \lambda_2 = 0.04$, $\lambda_3 = 0$, $\lambda_4 = 0.8$, $\frac{\delta}{\gamma} = 19$)

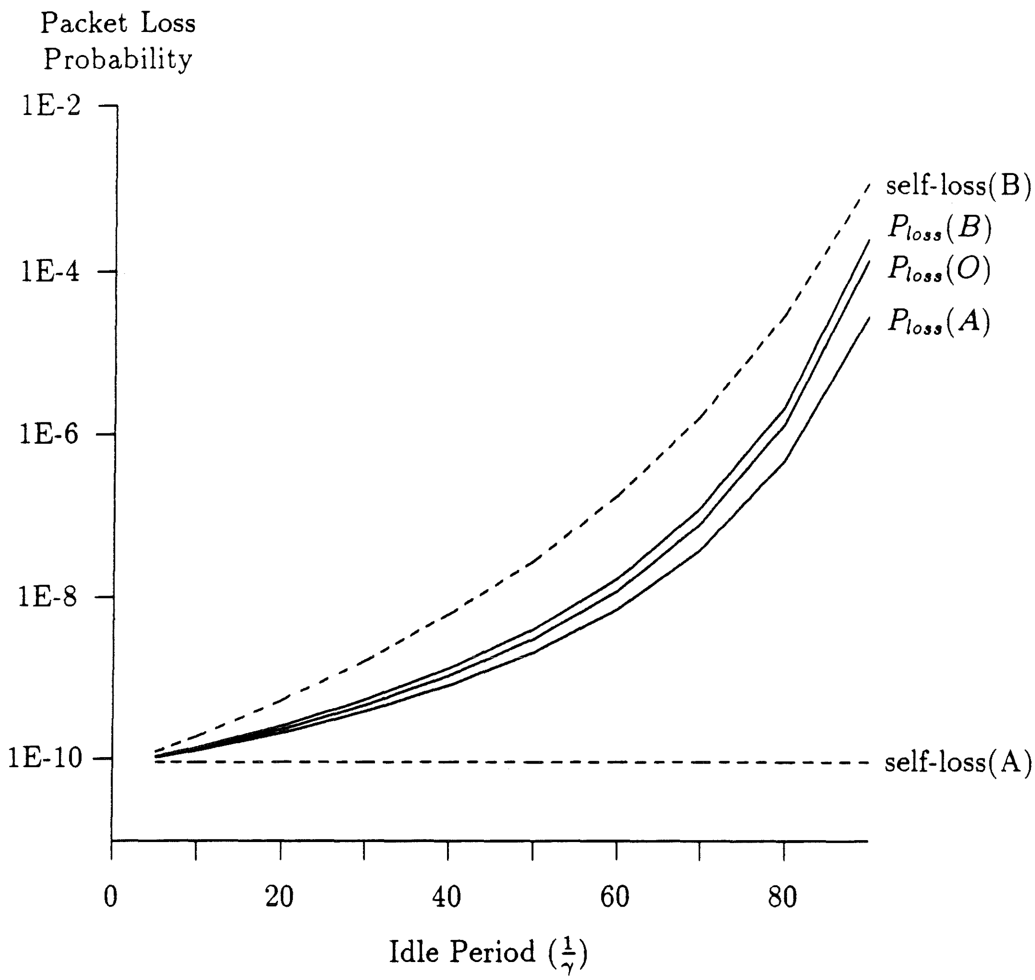

Figure 8. Effect of Burstiness on Packet Loss Probabilities Continuous Time, The Third Method

($\lambda_1 = \lambda_2 = 0.04$, $\lambda_3 = 0$, $\frac{1}{\gamma} + \frac{1}{\delta} = 100$)

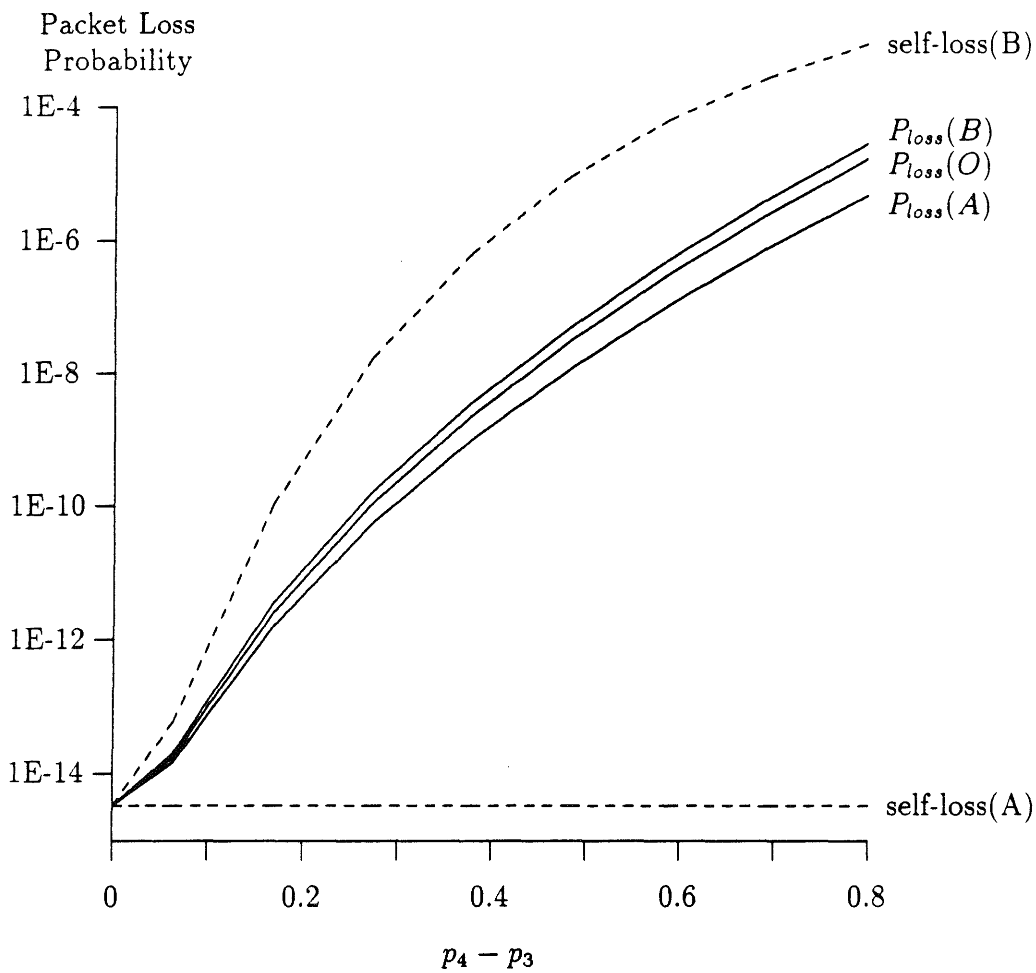

Figure 9. Effect of Burstiness on Packet Loss Probabilities
Discrete Time, The First Method
($p_1 = p_2 = 0.04$, $\gamma = 0.01$, $\delta = 0.19$)

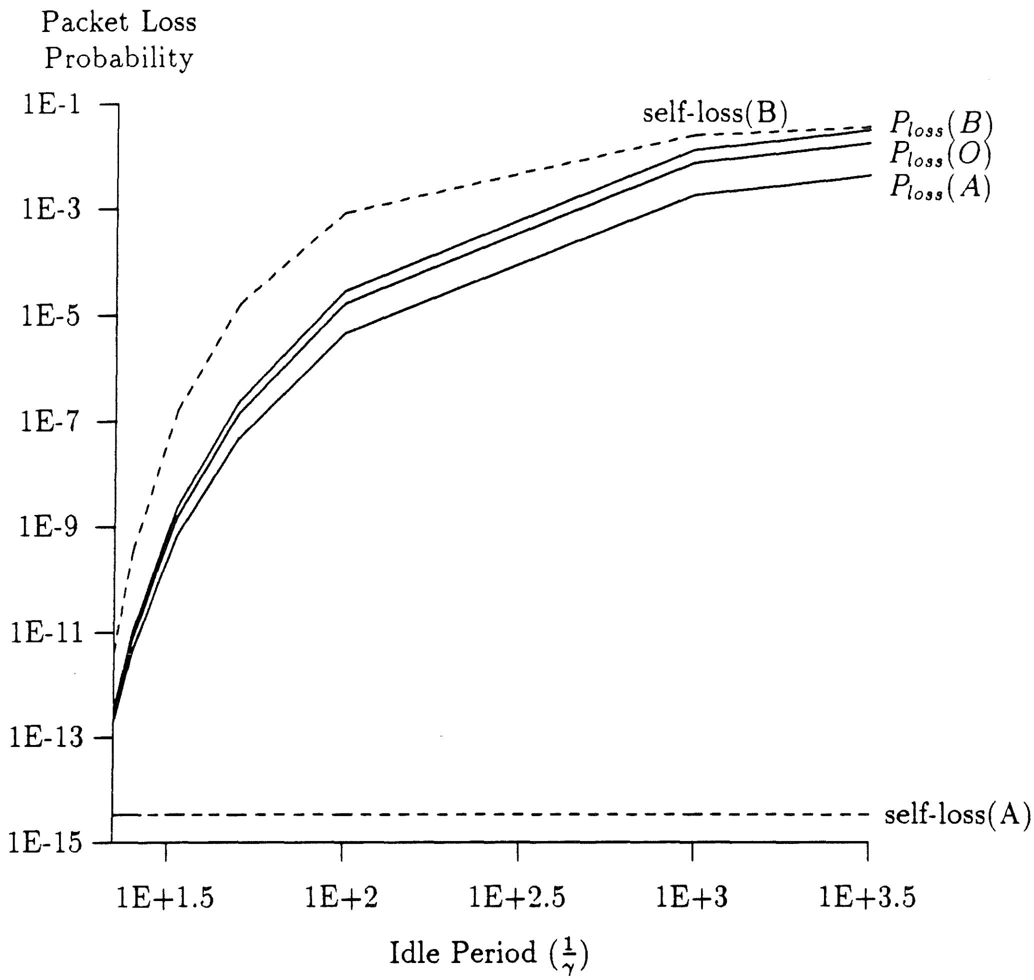

Figure 10. Effect of Burstiness on Packet Loss Probabilities Discrete Time, The Second Method

$(p_1 = p_2 = 0.04,\ p_3 = 0,\ p_4 = 0.8,\ \frac{\delta}{\gamma} = 19)$

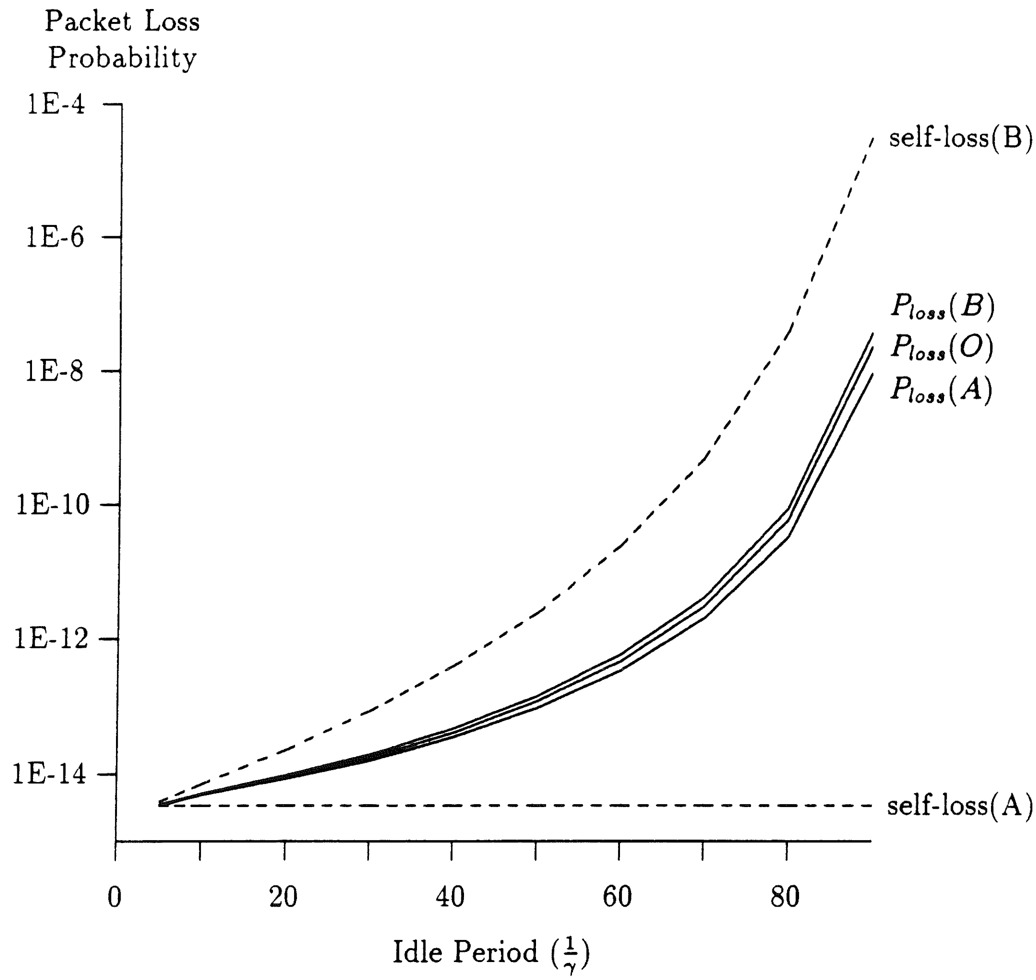

Figure 11. Effect of Burstiness on Packet Loss Probabilities
Discrete Time, The Third Method
($p_1 = p_2 = 0.04$, $p_3 = 0$, $\frac{1}{\gamma} + \frac{1}{\delta} = 100$)

$$
\begin{array}{c|ccccccccccc}
 & 0 & 1 & 2 & \cdots & N & N+1 & \cdots & K-2 & K-1 & K \\
\hline
0 & \{P_{1a}P_{1d}+P_{0a}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & \{P_{2a}P_{0d}+P_{3a}P_{1d}\}B & \cdots & P_{Na}P_{0d}B & & & & & \\
1 & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & \{P_{1a}P_{0d}+P_{2a}P_{1d}\}B & & \cdots & P_{Na}P_{0d}B & & & & \\
2 & & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P_{1a}P_{1d}\}B & & & \cdots & & & & \\
\cdot & & & & & & & & & & \\
\cdot & & & & & \vdots & & & & & \\
\cdot & & & & & & & & & & \\
K-N+1 & & & & & \cdots & & & \{P_{(N-3)a}P_{0d}+P_{(N-2)a}P_{1d}\}B & \{P_{(N-2)a}P_{0d}+P_{(\geq(N-1)a)}P_{1d}\}B & P_{(\geq(N-1)a)}P_{0d}B \\
K-N+2 & & & & & & \cdots & & \{P_{(N-4)a}P_{0d}+P_{(N-3)a}P_{1d}\}B & \{P_{(N-3)a}P_{0d}+P_{(\geq(N-2)a)}P_{1d}\}B & P_{(\geq(N-2)a)}P_{0d}B \\
\cdot & & & & & & & & & & \\
\cdot & & & & & \vdots & & & & & \\
\cdot & & & & & & & & & & \\
K-1 & & & & & & & & P_{0a}P_{1d}B & \{P_{0a}P_{0d}+P'_{a}P_{1d}\}B & P'_{a}P_{0d}B \\
K & & & & & & & & & P_{1d}B & P_{0d}B
\end{array}
$$

Figure 12. Transition Probability Matrix P for Multiple Stream Case (Discrete Time)